GEoREST Workshop on Induced Seismicity, 12th March 2024

Forecasting induced earthquake magnitudes using extreme value theory

James Verdon[1], Leo Eisner[2]

1. School of Earth Sciences
University of Bristol, UK.
`james.verdon@bristol.ac.uk`

2. Seismik s.r.o.
Prague, CZ

Outline

- **Objectives:** Forecasting M_{MAX}: why do we want to do this?
- **Methods:** Extreme value estimators
- **Test datasets:** Induced seismicity in Oklahoma, West Texas, and sequences from around the world ($n \geq 80$).
- **Performance metrics:** How do we define how well a model is performing?
- **Results**
- **Towards an empirically-constrained approach:** combining models to produce a strategy based on observed performance.

Motivation

Our objective is to forecast M_{MAX}, the largest event that is expected to occur during (or after) operations at a given site

- M_{MAX} (the largest event) controls the seismic hazard associated with a given operation.
- Maximum magnitude is of particular importance for regulators…

From: Department for Business, Energy & Industrial Strategy, Oil and Gas Authority, The Rt Hon Kwasi Kwarteng MP, and The Rt Hon Andrea Leadsom MP

Published 2 November 2019

- Oil and Gas Authority report published today concludes that it is not possible with current technology to accurately predict the probability of tremors associated with fracking
- Separate proposals to change the planning process for fracking sites will no longer be taken forward at this time

Fracking will not be allowed to proceed in England, the government has announced today, following the publication of new scientific analysis.

Motivation

Need for transparent, robust and extensive testing:

- Build confidence in model performance with regulators/operators
- Establish whether models can be generalized well (or if not, what types of situation lead to better or worse model performance)
- Identify common (or different) behaviours/trends between different modelling methods, perhaps leading to a more holistic strategy
- Performance assessment can feed back into physical understanding of induced seismicity

Methods

A range of methods have been used to forecast induced seismicity magnitudes:

- Numerical geomechanical simulations…

- Statistics-based methods – quick to parameterize and operate in real time:
    - Scaling between seismicity rates and injection parameters…
    - Machine learning…
    - Extreme value theory…

Methods

Upper limit and record-breaking event theory (Cooke, 1979).

- Applied in PSHA for tectonic earthquakes (e.g., Kijko, 2004)
- Applied to mining-induced seismicity by Mendecki (2016)

M_{UL}:

$$M_{UL} = 2M_n^O - \sum_{i=1}^{n-1}\left[\left(1-\frac{i}{n}\right)^n - \left(1-\frac{i+1}{n}\right)^n\right] M_{n-i}^O$$

- $M^o{}_i$ are the observed event magnitudes, ordered from smallest to largest (so $M^o{}_n$ is the largest observed event magnitude).
- This theory is not specific to earthquake magnitudes. We can alternatively use earthquake moment or potency ($P = M_O / \mu$) instead.
- Hereafter, we use M_{UL_MM} for magnitude-based calculations, and M_{UL_MO} when using potencies

Methods

Upper limit and record-breaking event theory (Cooke, 1979).

- Alternatively, we can estimate the maximum expected magnitude jump (or increment).
- We refer to this estimate as the "jump-limited" case:

***M_{JL}*:**

$$M_{JL} = M^{O}_{MAX} + \Delta M_{MAX}$$

$$\Delta M_{MAX} = 2\Delta M^{O}_{n_j} - \sum_{i=1}^{n_j-1}\left[\left(1-\frac{i}{n_j}\right)^{n_j} - \left(1+\frac{i+1}{n_j}\right)^{n_j}\right]\Delta M^{O}_{n_j-i}$$

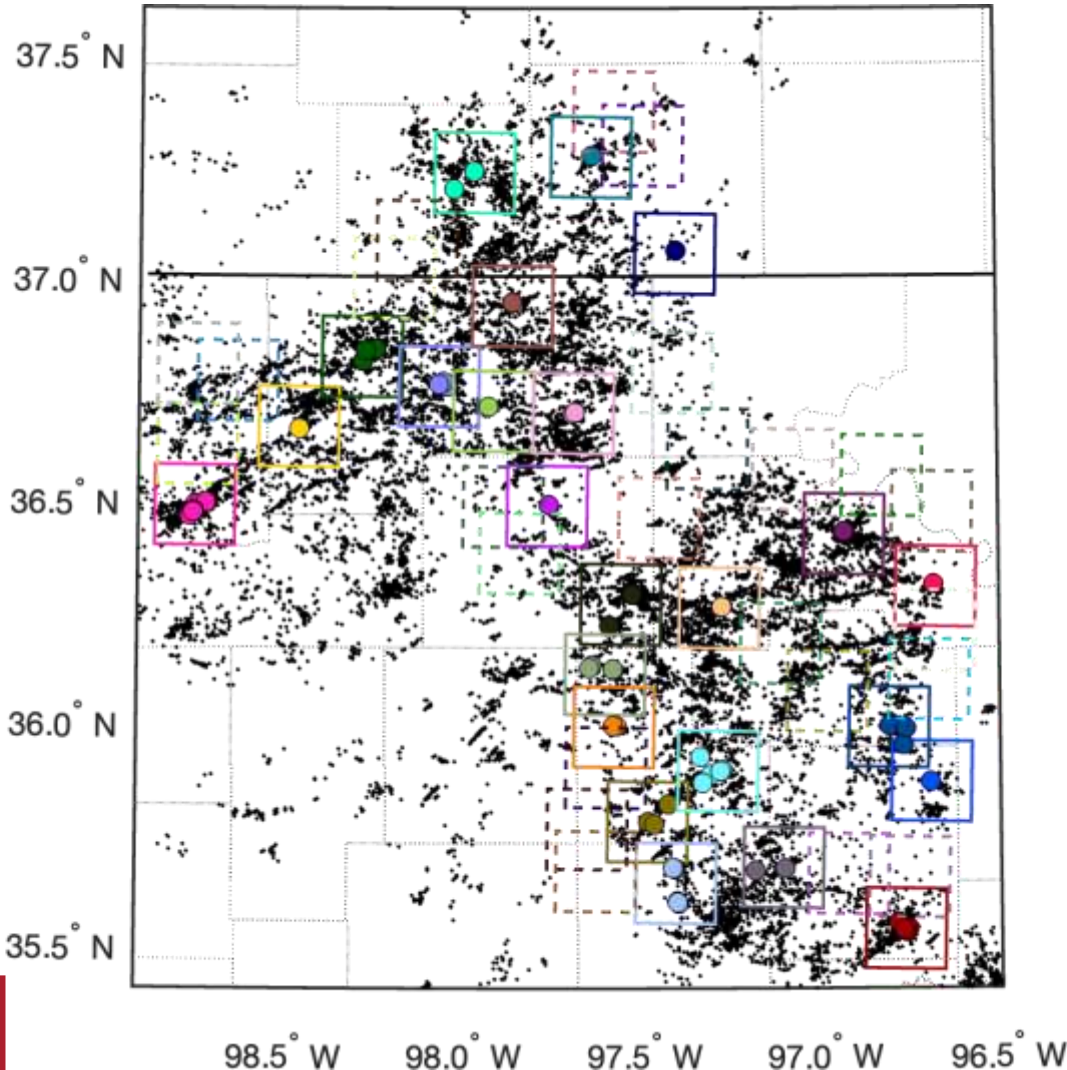
- ΔM^{o}_{i} are the observed magnitudes jumps (increases over all previous events), ordered from smallest to largest (so ΔM^{o}_{n} is the largest observed event magnitude jump).

Datasets

Oklahoma

- Event catalogue from Park et al. (2022) using PhaseNet: $M_C \approx 1.0$.

- 20 x 20 km boxes around all sequences around M4+ events [24 cases].
- An additional 24 boxes mapped at random which contain at least 500 events but no M 3.5 or above.
- Because we want to test how the models perform with sequences that don't generate large events.

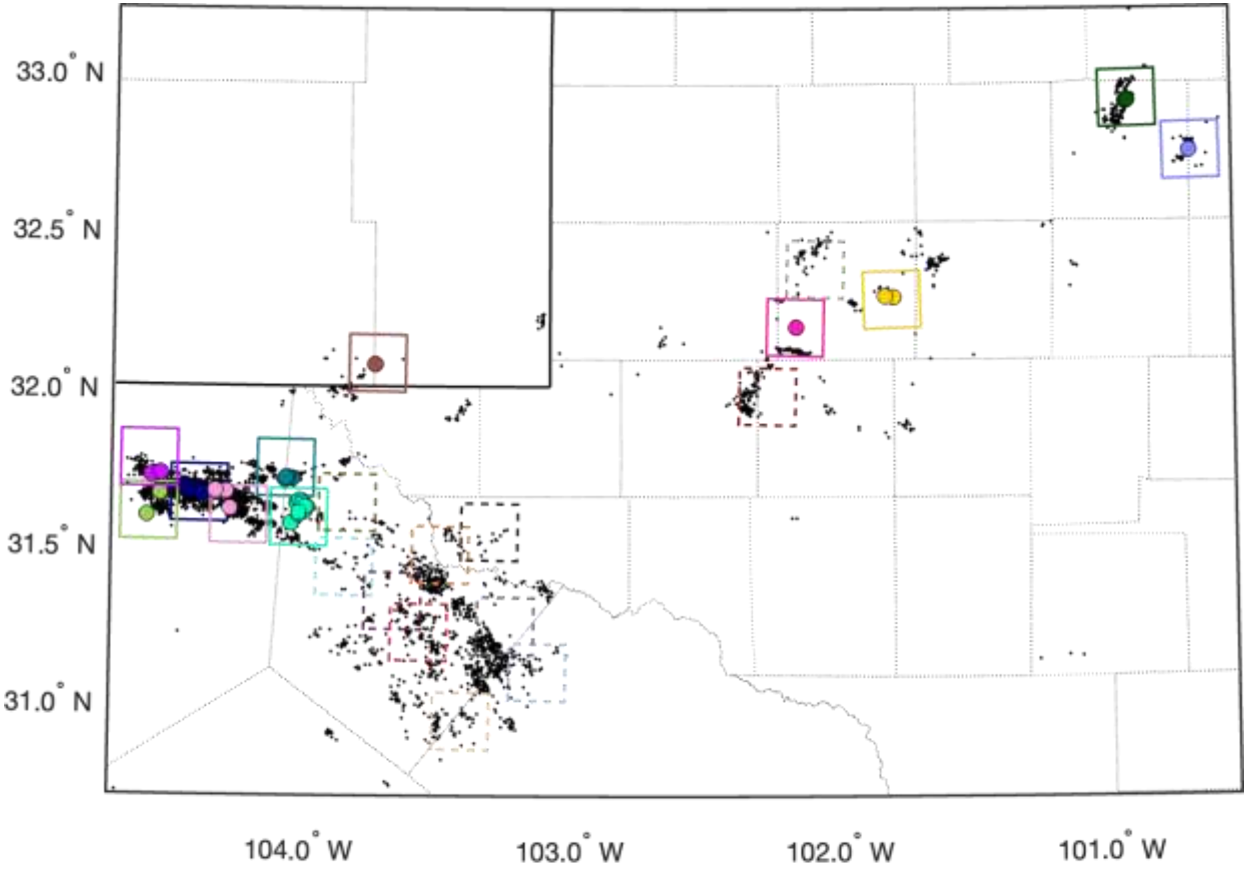


Datasets

West Texas

- Event catalogue from TexNet: $M_C \approx 2.0$.

- 20 x 20 km boxes around all sequences around M 4+ events [11 cases]
- An additional 11 boxes mapped at random which contain at least 100 events but no M 3.5 or above.



Datasets

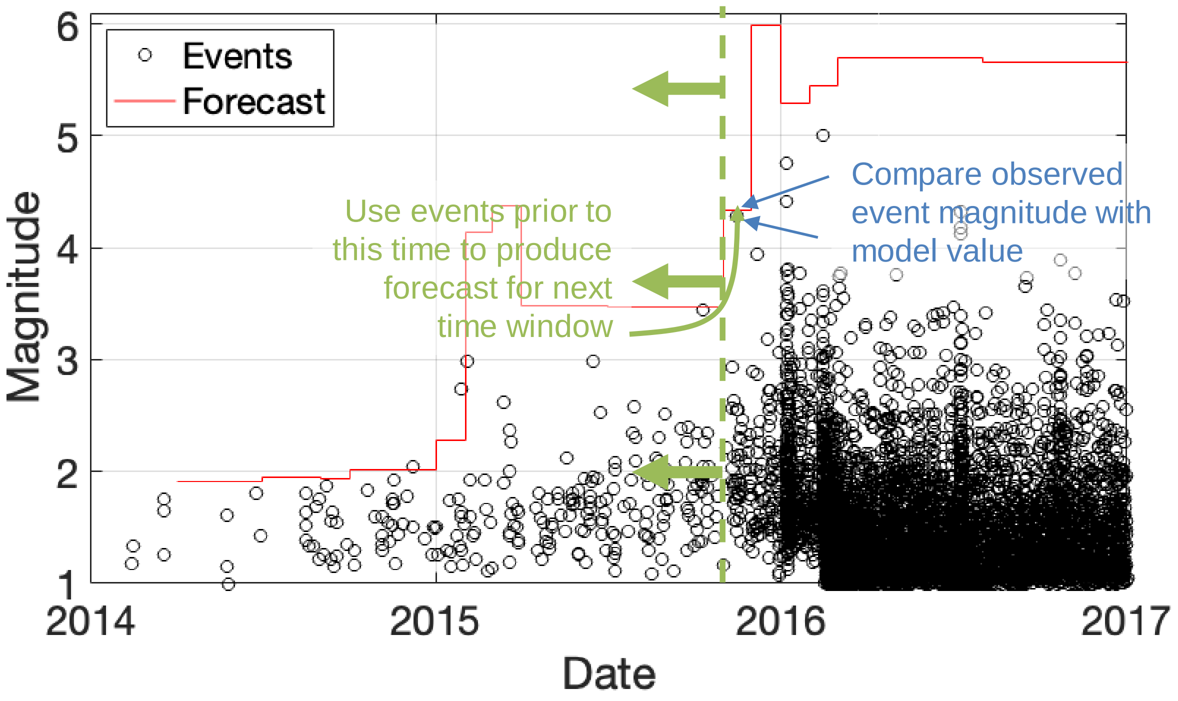
From Watkins et al. (2023)

- 16 cases: eastern Texas (Azle-Reno, Dallas-FW, Irving, Timpson, Venus, Guy-Greenbrier, Youngstown, Paradox, Greeley, Raton Basin, Eagle West, Graham, Musreau Lakes, Rongchang, Castor, Rubiales.

Model Performance

We make forecasts in a pseudo-prospective manner:

- At a given time, all previous events are used to estimate M_{MAX}
- If the next window contains a new "record-breaking" event (M^{O}_{MAX}) then we compare the observed event magnitude with the modelled values



Model Performance

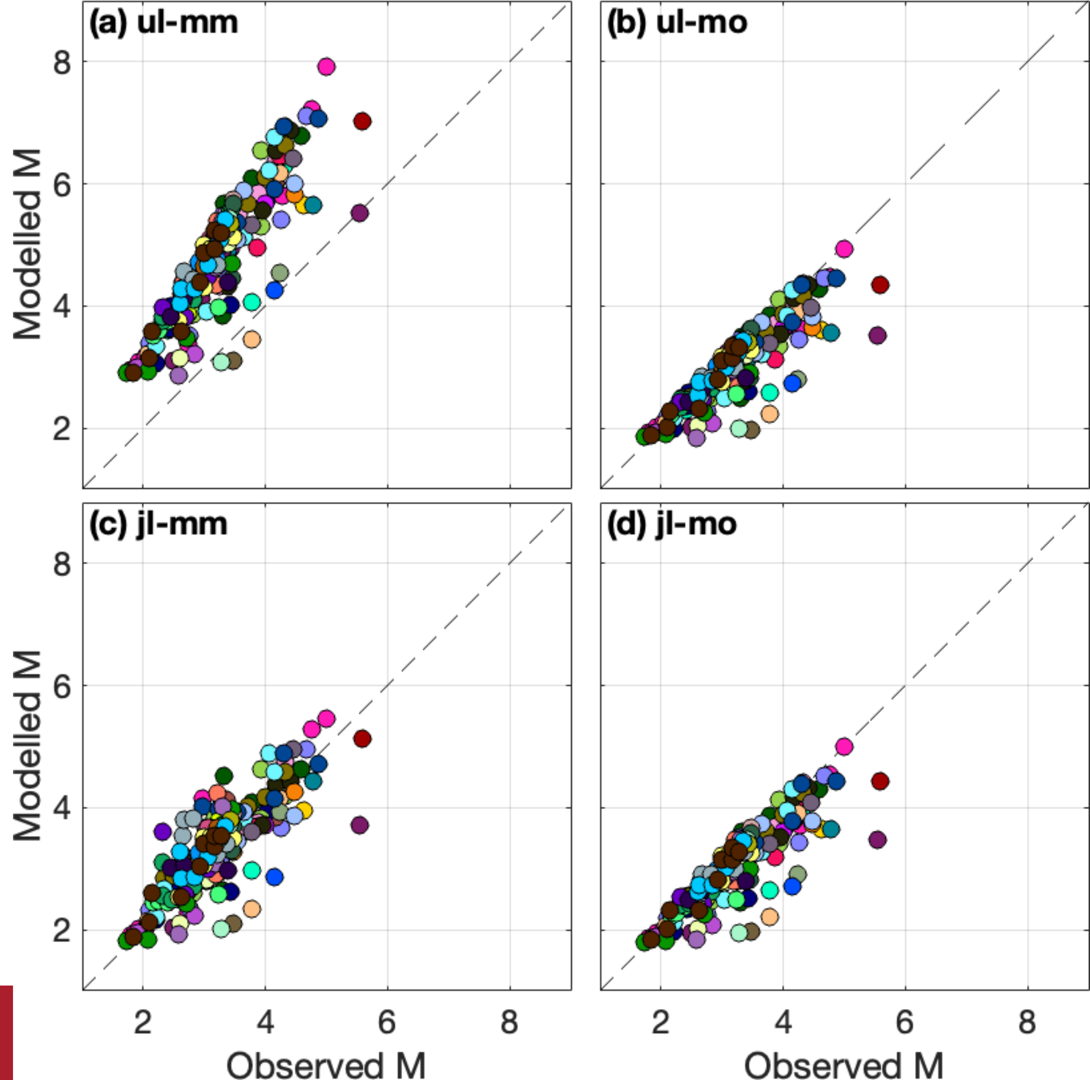
- In total we have **86 individual sequences**, in which a combined total of **331 record breaking** events occur. Very large dataset for a comprehensive assessment of forecasting performance.

Performance Metrics:

- **RMS error, σ_{RMS}:** how close are modelled and observed mags?
- **Pearson correlation coefficient, r:** do the modelled and observed mags fall along a line (not necessarily 1:1), or are they randomly scattered?
- **Line of best fit between modelled and observed mags, m:** does the best fit line fall along a 1:1 ratio?
- **How many significant underpredictions, N_{UP}:** cases where the observed magnitude is more than 0.5 units above the forecast value. We want forecasting models that do not underestimate the risk, since we want to make conservative decisions.

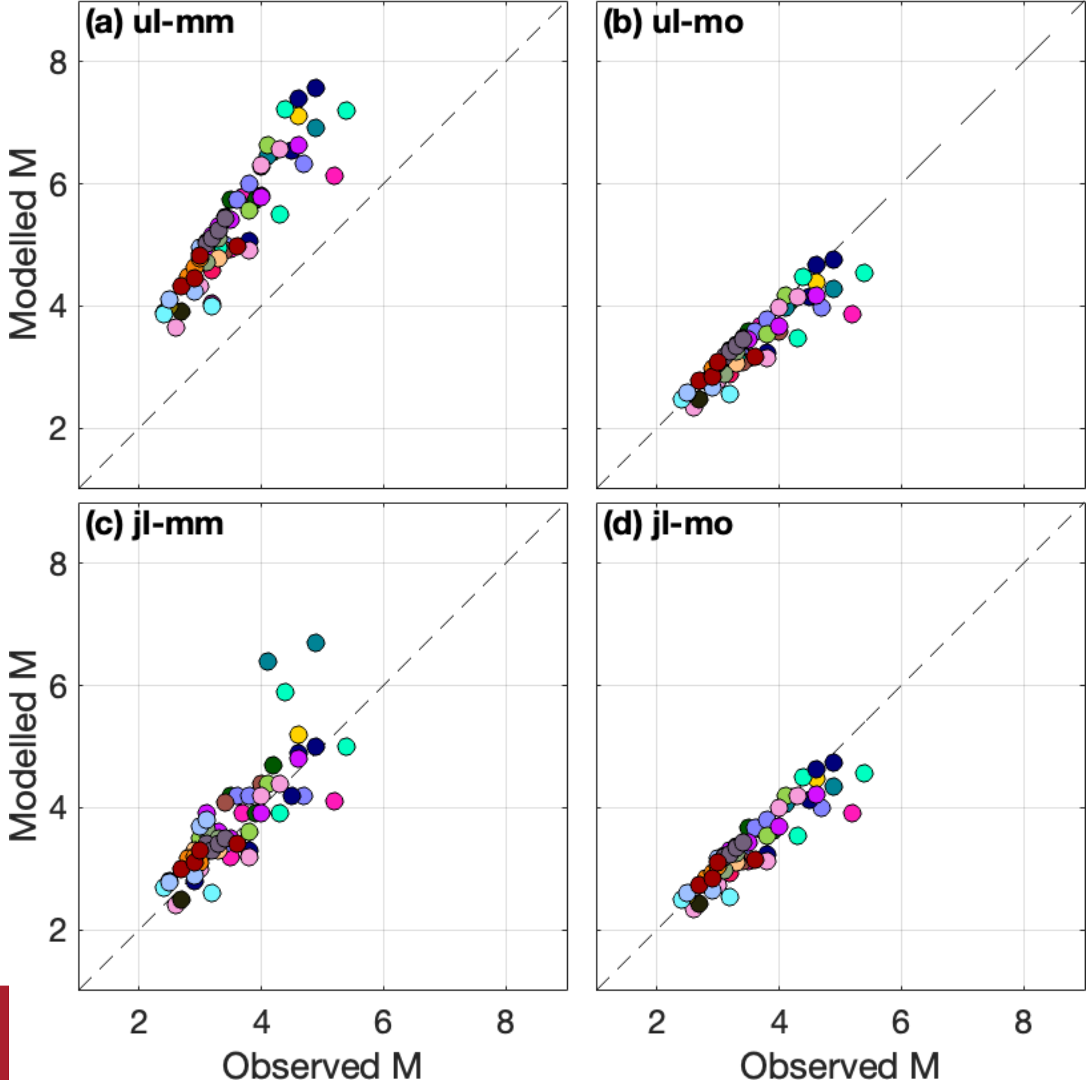
Results

- Results for OK/KS:



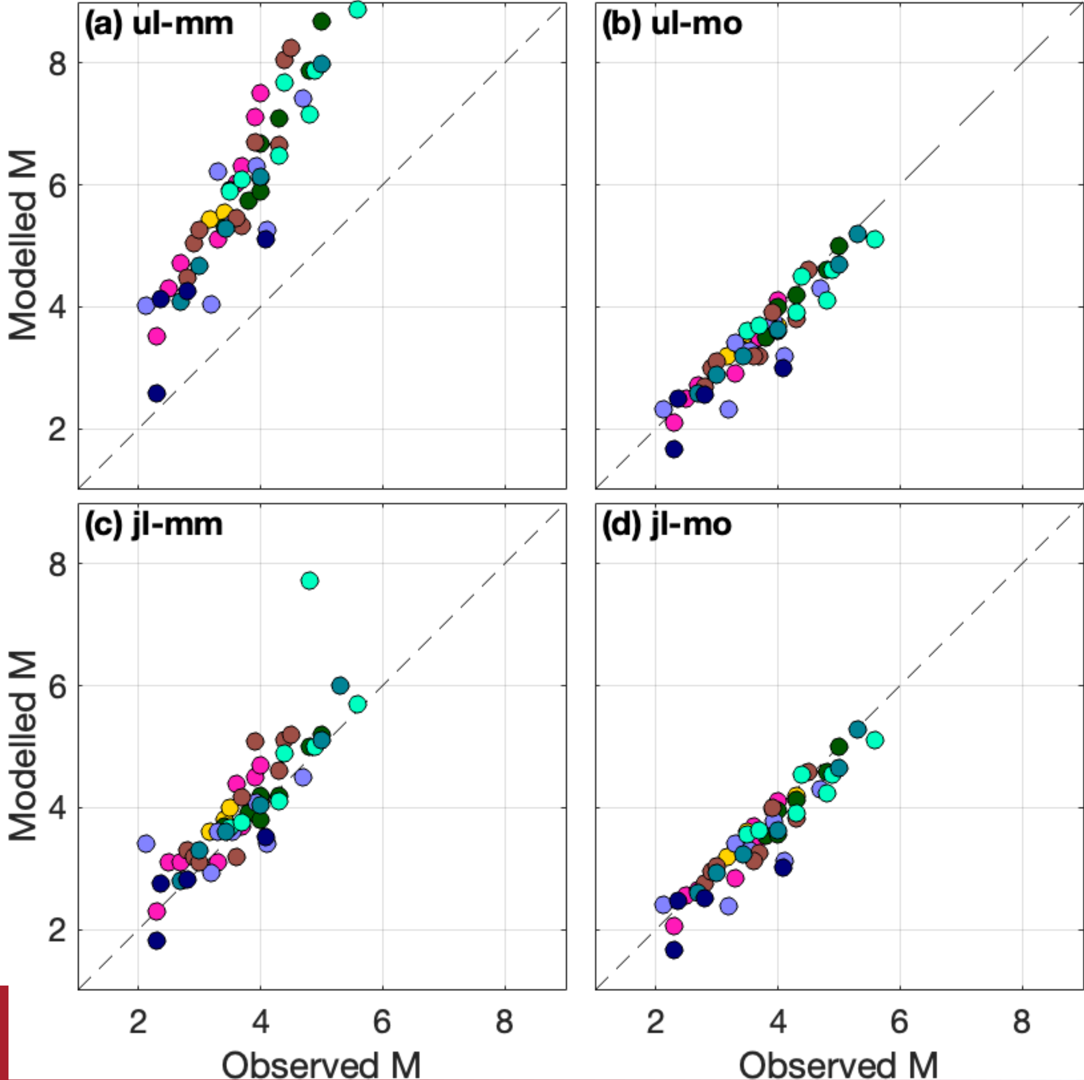
Results

- Results for west Texas:



Results

- Results for Watkins et al. sequences

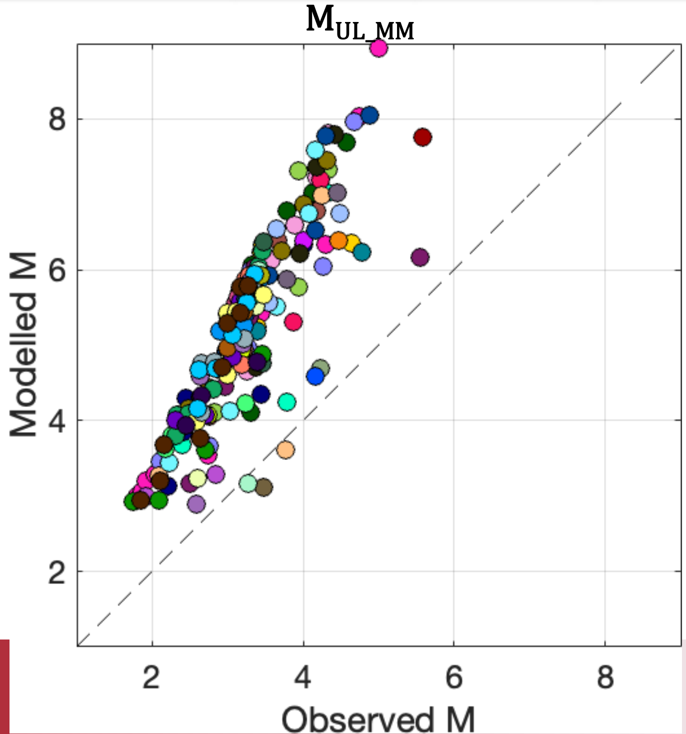


Results

Upper Limit values using magnitudes

- Generally over-predict.
- Largest misfit between observed and modelled values.
- *m* >> 1 – line is not 1:1.
- But correlation is good – it is a line, just not where we want it to be.
- Never under-predicted. A credible upper limit that won't be exceeded.

Model	σ_{RMS}	r	m	N_{UP} [%]
M_{UL_MM}	1.84	0.86	1.27	0
M_{UL_MO}	0.41	0.86	0.76	14.2
M_{JL_MM}	0.47	0.81	0.85	7.3
M_{JL_MO}	0.41	0.85	0.78	14.6

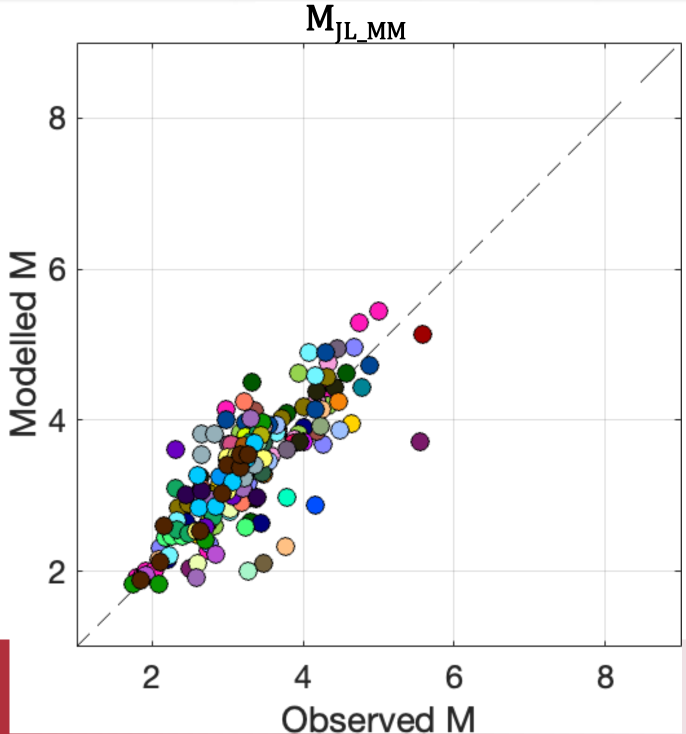


Results

Jump-Limited values using magnitudes

- Fit the shape of the data relatively well ($m \sim 1$).
- However, they have the largest scatter – second largest RMS error, and smallest correlation coefficients.
- Seldom create underpredictions

Model	σ_{RMS}	r	m	N_{UP} [%]
M_{UL_MM}	1.84	0.86	1.27	0
M_{UL_MO}	0.41	0.86	0.76	14.2
M_{JL_MM}	0.47	0.81	0.85	7.3
M_{JL_MO}	0.41	0.85	0.78	14.6

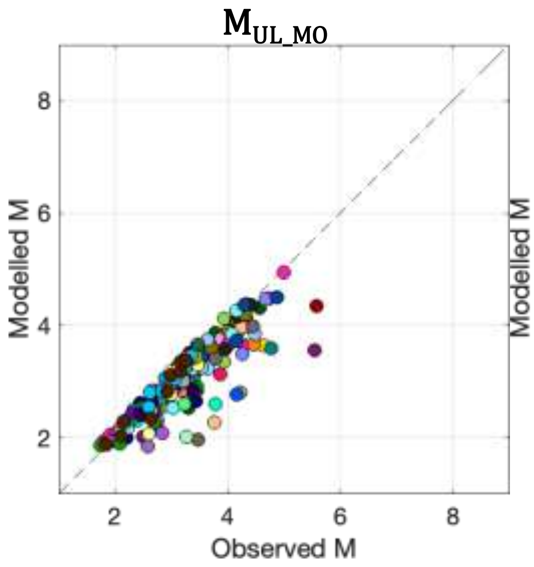


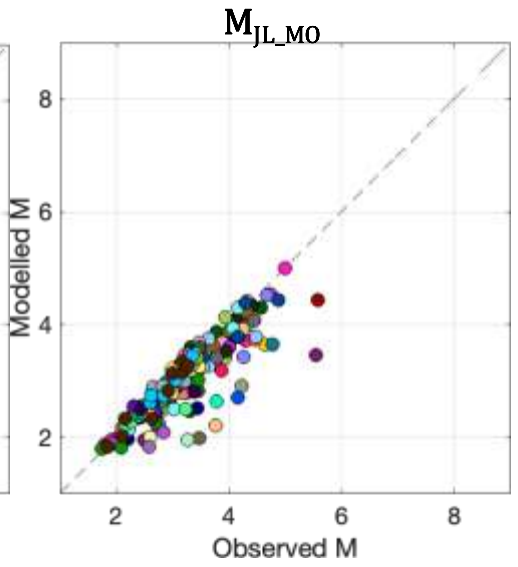
Results

Potency-based methods

- Provide the best fit to the data (lowest RMS errors).
- Provide the least scatter (highest correlation coefficient)
- On occasion, produce significant under-predictions, which could be an issue.
- While most results fit along a 1:1 line, the handful of underpredictions produce a lower best-fit gradient.

Model	σ_{RMS}	r	m	N_{UP} [%]
M_{UL_MM}	1.84	0.86	1.27	0
M_{UL_MO}	0.41	0.86	0.76	14.2
M_{JL_MM}	0.47	0.81	0.85	7.3
M_{JL_MO}	0.41	0.85	0.78	14.6





Towards an empirically-constrained model

- A large number of cases provides a robust and comprehensive testing dataset.
- We can empirically constrain how each method performs. This constraint can then be used to adapt our model (or suite of models) to produce a more useable forecast with more realistic expectations.
- For example, if we had some models that always over-predict, and some models that always under-predict, then we can assume that the true magnitude will be somewhere between these bounds…

Towards an empirically-constrained model

- Potency-based models usually provide a good fit to the data, but occasionally under-predict.
- M_{UL} using magnitudes usually over-predicts, but never under-predicts.

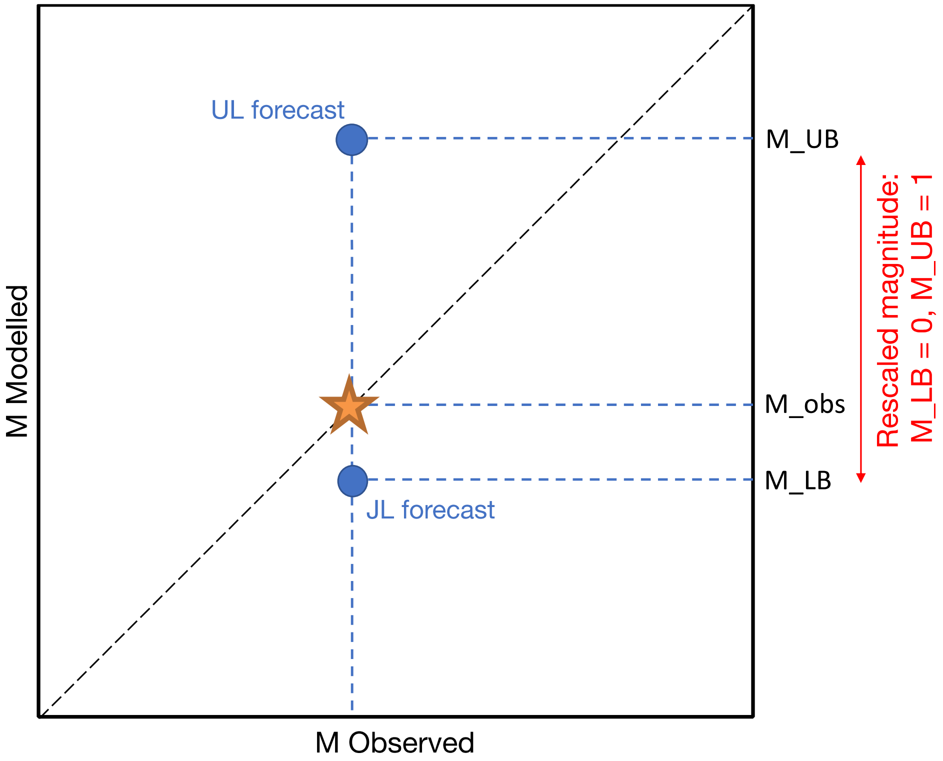
- Therefore, we should expect the actual M_{MAX} to fall somewhere between the M_{UL_MM} and M_{JL_MO} values. M_{UL_MM} is the upper bound, M_{UB}, and M_{JL_MO} is the lower bound, M_{LB}.

- How do observed magnitudes distribute between these upper and lower bounds?

Towards an empirically-constrained model

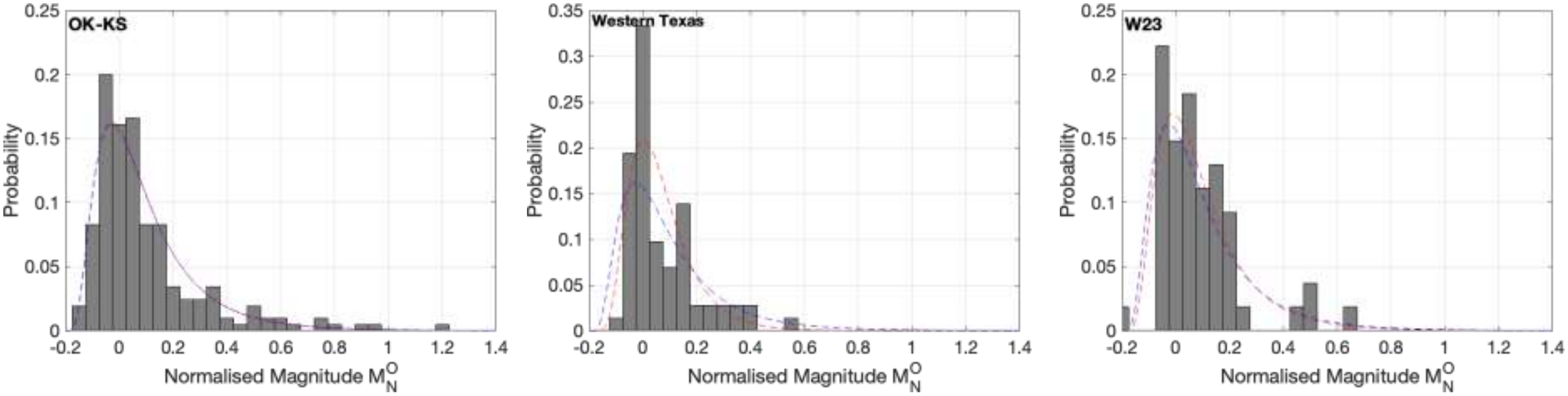
- We normalise each observed $M^O{}_{MAX}$ event by the M_{UB} and M_{LB} values at the time the event occurred:

$$M_N^O = \frac{M_{MAX}^O - M_{LB}}{M_{UB} - M_{LB}}$$



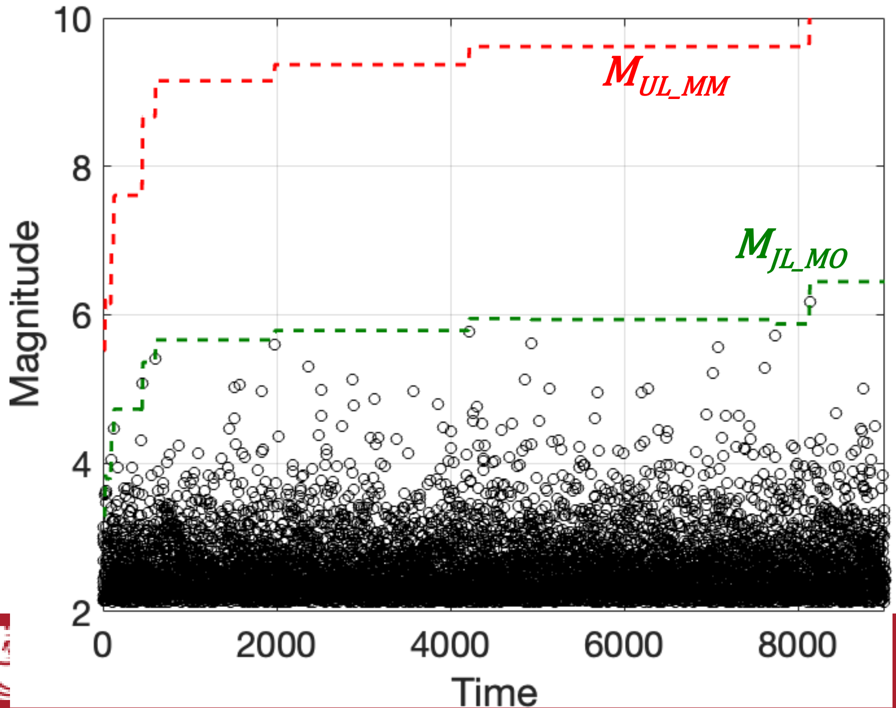
Towards an empirically-constrained model

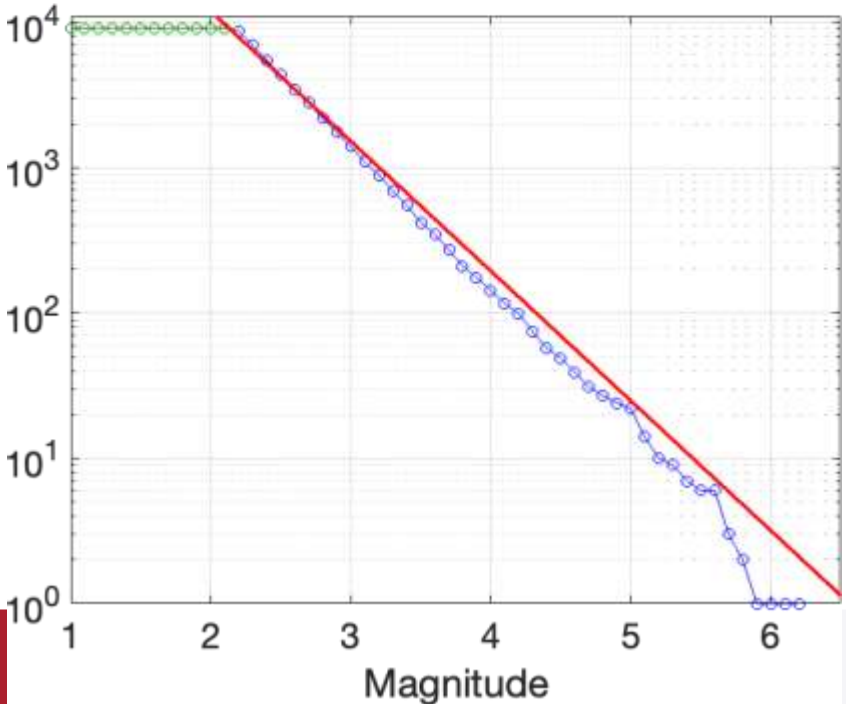
- Behaviour is really consistent between study areas – most cases clustered around 0 (i.e., $M^{OBS}{}_{MAX} \approx M_{LB}$), but a tail of events reaching towards 1 (i.e., $M^{OBS}{}_{MAX} = M_{UB}$).
- Shifted lognormal distributions seem to fit the data well (red curves): μ_{LN} = [- 1.4, -1.41, -1.37], σ_{LN} = [0.59, 0.43, 0.54], with a shift of δ = -0.2.



Towards an empirically-constrained model

- Testing using synthetic catalogs. We generate 1,000 random catalogs.
- *N* chosen randomly each time (between 500 – 10,000). M_{MIN} chosen randomly each time (between 0.5 – 2.5). *b* is always 1.0. Generate random events from an unbounded G-R distribution.

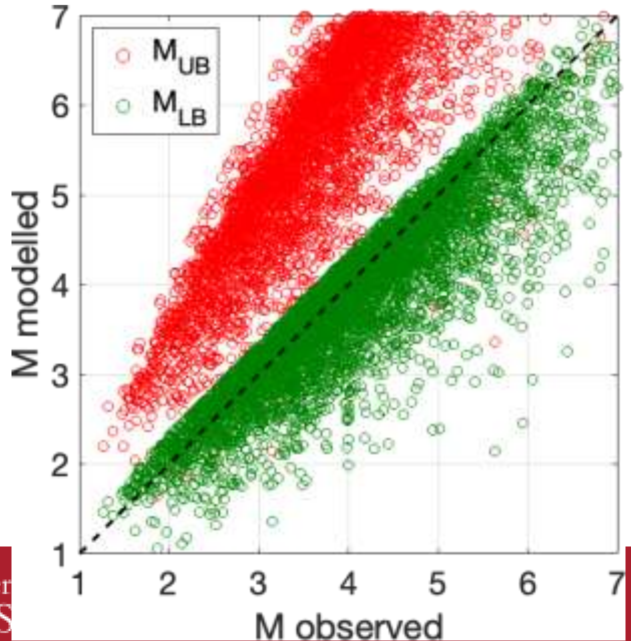


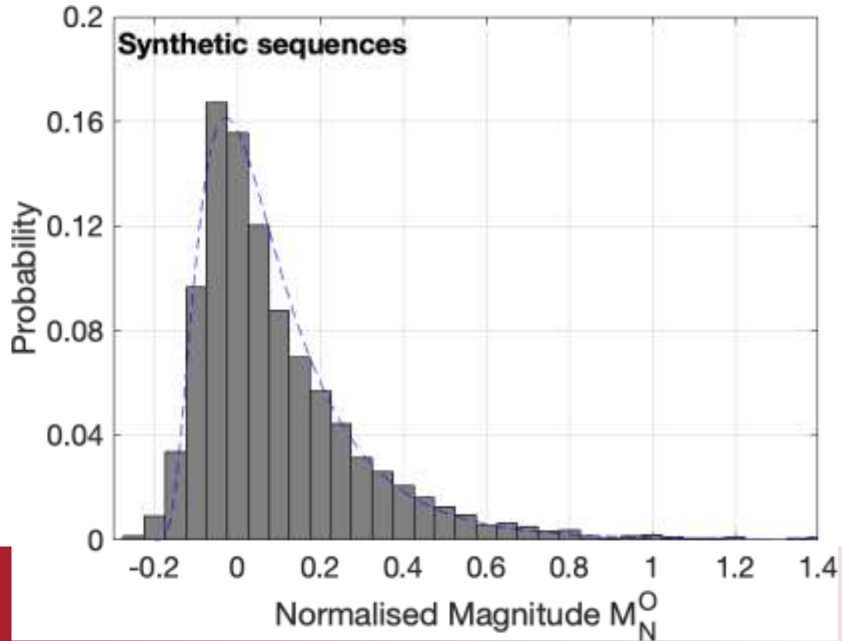


Towards an empirically-constrained model

- Results look a lot like our observations – M_{UB} overpredicts but is a credible upper limit. M_{LB} generally fits the data well, but occasionally underpredicts.

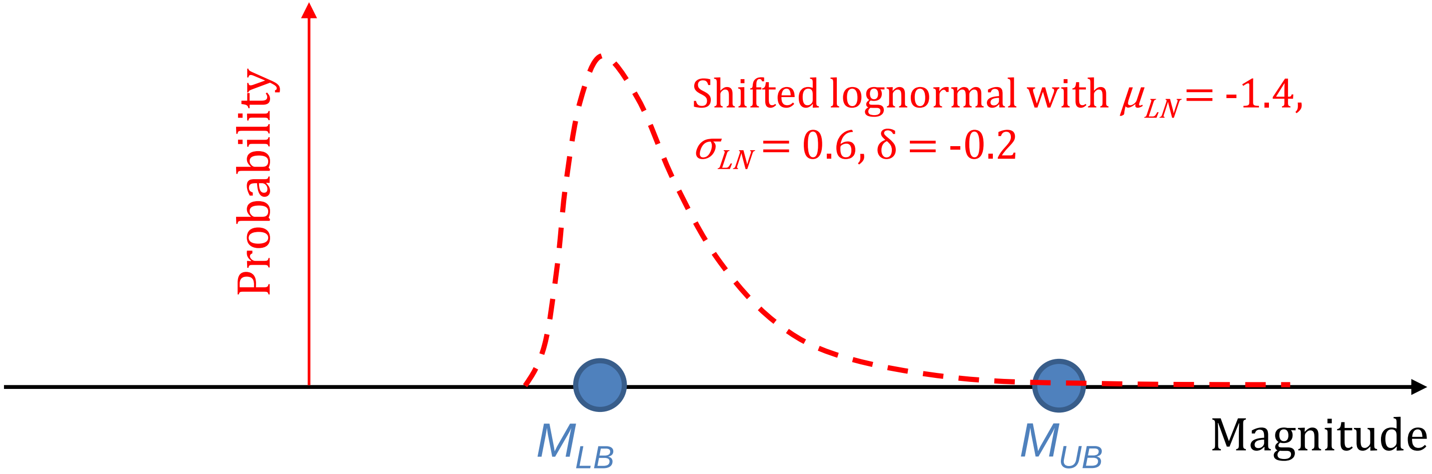
- Normalised magnitudes are well fit by a shifted lognormal distribution, with very similar values to our observed cases: μ_{LN} = -1.4, σ_{LN} = 0.6, δ = -0.2.





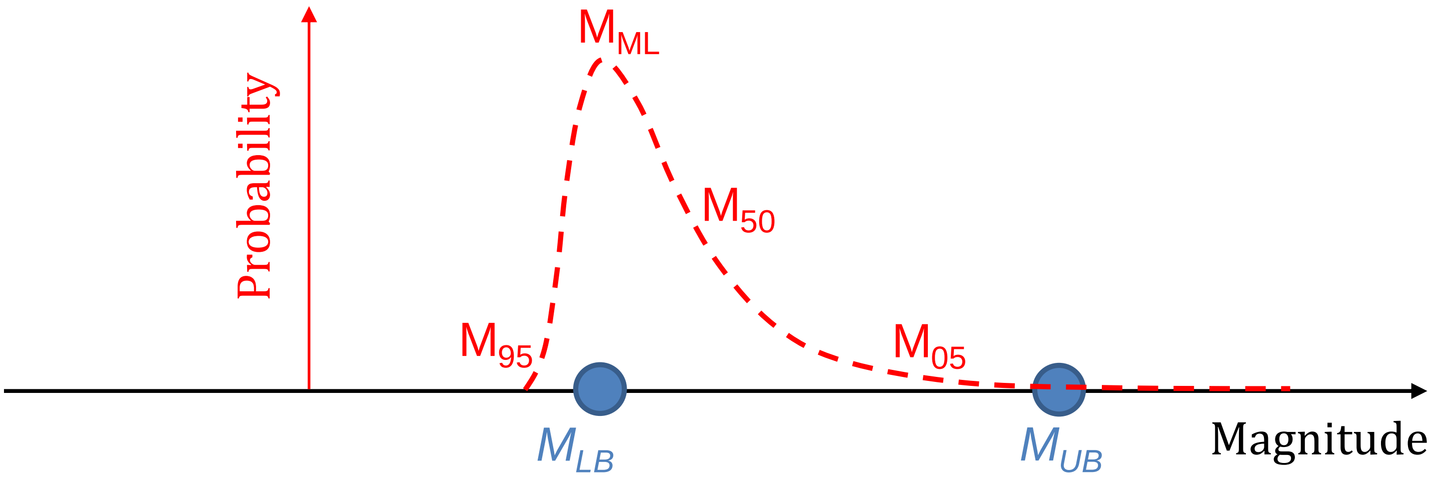
Towards an empirically-constrained model

- It looks like this distribution is an inherent result when sampling magnitudes from an underlying G-R.
- We can use this behaviour to produce a probabilistic estimate for the next record-breaking event during a sequence.
- Compute M_{LB} *and* M_{UB}, and assign probability values to magnitudes between these values:



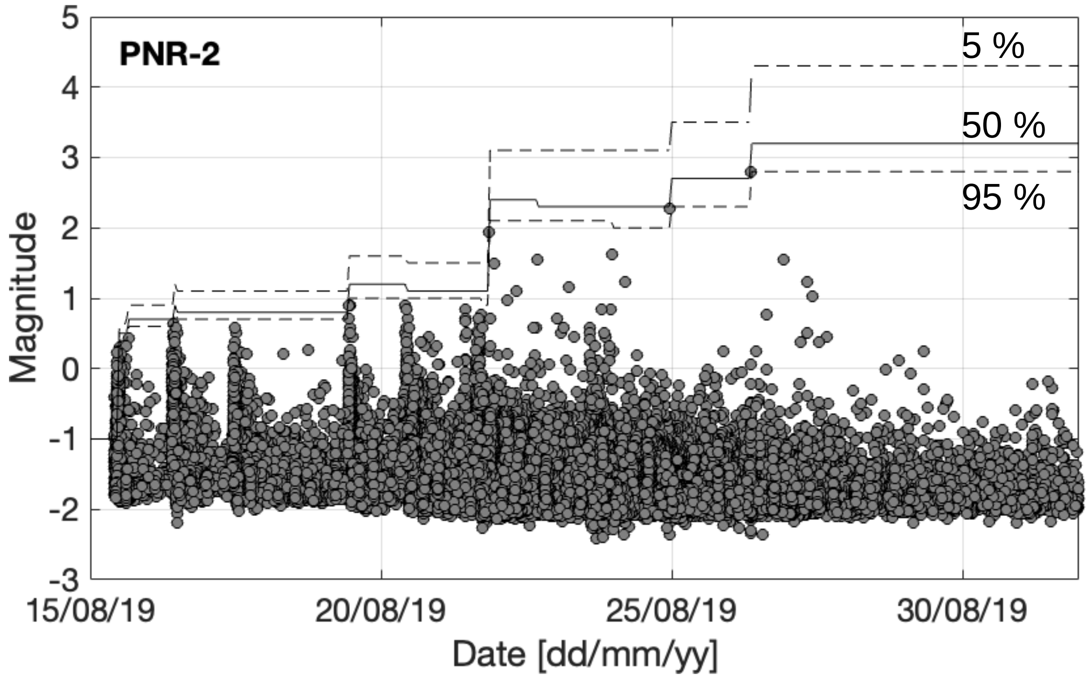
Towards an empirically-constrained model

- It looks like this distribution is an inherent result when sampling magnitudes from an underlying G-R.
- We can use this behaviour to produce a probabilistic estimate for the next record-breaking event during a sequence.
- Compute M_{LB} *and* M_{UB}, and assign probability values to magnitudes between these values:



Towards an empirically-constrained model

- **Demo:** Application to PNR-2 – an out of sample case that was not used to define the probability distribution
- The M 2.8 is well within the forecast range, close to the M_{50} value.
- The M 1.9 event after Stage 6 was not well forecast ($p = 1$ %).
- Stage 6 has been identified as the time the microseismicity changed – switch from HF propagation to fault reactivation.
- Not surprising that HF microseismicity is not good for forecasting fault reactivation.



Conclusions

- M_{MAX} forecasting using extreme value statistics shows significant potential
- We have tested various implementations of this approach across a very large number of case studies. We find good correlation between forecast and observed maximum magnitudes
- There are systematic differences in the results produced by different implementations. These differences are found in both observed and synthetic datasets.
- We use these differences to define an empirically-constrained M_{MAX} estimation:
    1. Upper-limit estimation based on magnitudes defines the upper bound
    2. Jump-limit estimation based on potency defines the “lower” bound
    3. Likelihood is estimated from a distribution between these bounds

Acknowledgements

Bristol and Oxford Passive Seismic Research Consortium

seisgreen

UKRI Natural Environment Research Council

Thanks!

Any questions, comments or suggestions?

james.verdon@bristol.ac.uk

Bristol and Oxford Passive Seismic Research Consortium